Figure 2

The empiric cumulative and the distribution $F(\hat{\alpha}, \hat{\beta})$ for fatigue life at a stress of 26,000 psi.

ESTIMATION FOR A FAMILY OF LIFE DISTRIBUTIONS

WITH APPLICATIONS TO FATIGUE

by

Sam C. Saunders

and

Z. W. Birnbaum
University of Washington
Seattle, Washington

Summary

The estimation problem is studied for a new two-parameter family of life length distributions which has been previously derived from a model of fatigue crack growth. Maximum likelihood estimates of both parameters are obtained and iterative computing procedures are given and examined. A simple estimate of the median life is exhibited, shown to be consistent and then compared, favorably, with the maximum likelihood estimate. More importantly the asymptotic distribution of this estimate is shown to be within the same class of distributions as the observations themselves. This model, and these estimation procedures, are tried by fitting this distribution to several extensive sets of fatigue data and then some comparisons of practical significance are made.

1. Introduction

A new family of distributions was derived in [1] from considerations of the physical behavior of fatigue crack growth under cyclic loading. The life length obtained was the mathematical representation of the number of cycles needed to force the fatigue crack to exceed a critical value.

Let us denote by $\mathcal{G}$ this two-dimensional parametric family of distributions of nonnegative random variables defined by

$$F(t:\alpha,\beta) = \mathfrak{N}[\frac{1}{\alpha} \xi(t/\beta)] \quad \text{for} \quad t > 0 \tag{1.1}$$

where $\alpha > 0$, $\beta > 0$ and

$$\xi(t) = t^{\frac{1}{2}} - t^{-\frac{1}{2}} \tag{1.2}$$

and $\mathfrak{N}$ is the distribution function of the standard normal variate.

Of course, there have been many families of distributions which have been suggested as candidates for the theoretical probability law governing fatigue life. Excellent discussions and comparisons of many of these have been made by other authors, see for example [3], [5], and [6] and the references given there. However, the main intent of this study is to investigate the parametric estimation problem for the family defined in (1.1) and not to argue its particular merits in such applications over other families of distributions.

In this paper, we derive the maximum likelihood estimates of the parameters α and β and develop some iterative numerical procedures for their computation. We obtain a simple estimate of β, which is

shown to be consistent and shown, for small values of α, to be virtually the same as the maximum likelihood estimate. In all cases, it can be used as a good initial guess for the iterative procedures which compute the maximum likelihood estimate. We also obtain the bias and variance of this simplified estimate under certain limiting conditions. Two different numerical procedures for computing the maximum likelihood estimate of β are proposed and their behavior compared by applying them to several large sets of observations with different, but known, values of the parameters. We finally present the results of fitting this distribution by these estimation procedures to some rather extensive sets of fatigue data obtained from metal coupons cycled at various stress levels.

In order to make this paper self-contained we now quote without proof some results given in [1], which will be used subsequently.

Theorem 1.1. If T has the life distribution $F(\alpha,\beta)$ in $\mathcal{G}$ then $1/T$ has the distribution $F(\alpha,\frac{1}{\beta})$ also in $\mathcal{G}$, and for any real $a > 0$ the random variable aT has a distribution $F(\alpha,a\beta)$ in $\mathcal{G}$. Moreover,

$$ET = \beta(1 + \frac{\alpha^2}{2}) \tag{1.3}$$

$$\text{var}(T) = (\alpha\beta)^2(1 + \frac{5\alpha^2}{4}) \tag{1.4}$$

and if Z is a standard normal variate, then

$$1 + \frac{\alpha^2}{2} z^2 + \alpha z\sqrt{1 + \frac{\alpha^2}{4} z^2} \tag{1.5}$$

has the distribution $F(\alpha,1)$ in $\mathcal{G}$.

2. Maximum Likelihood Estimates and Their Computation

For a given set of positive numbers $t_1,\ldots,t_n$ define the arithmetic and harmonic means by

$$s = \frac{1}{n} \sum_1^n t_i, \qquad r = \left[\frac{1}{n}\sum_1^n t_i^{-1}\right]^{-1} \tag{2.1.1}$$

and the harmonic mean function K by

$$K(x) = \left[\frac{1}{n} \sum_1^n (x+t_i)^{-1}\right]^{-1} \qquad \text{for } x > 0. \tag{2.1.2}$$

Now we can state the primary

<u>Theorem 2.1</u>. If $t_1,\ldots,t_n$ is a sample of independent observations each distributed by $\Gamma(\alpha,\beta) \in \mathcal{G}$, then the maximum likelihood estimate $\hat{\beta}$ of β is the unique positive solution of $g(x) = 0$ where g is the random function defined by

$$g(x) = x^2 - x[2r + K(x)] + r[s + K(x)]. \tag{2.1.3}$$

Furthermore $r < \hat{\beta} < s$. We can also express the maximum likelihood estimate $\hat{\alpha}$ of α, defined in terms of $\hat{\beta}$, by

$$\hat{\alpha} = \sqrt{\frac{s}{\hat{\beta}} + \frac{\hat{\beta}}{r} - 2}. \tag{2.1.4}$$

Proof. Consider the density for $t > 0$ obtained from (1.1)

$$F'(t:\alpha,\beta) = \frac{1}{\alpha\beta} \, \mathfrak{N}'[\frac{1}{\alpha} \, \xi(t/\beta)] \xi'(t/\beta).$$

Taking the natural logarithm of the joint density of the observations $t_1, \ldots, t_n$ we obtain the likelihood

$$L = -n \ln \alpha - n \ln \beta + \sum_{i=1}^{n} \{ -\tfrac{1}{2} \ln(2\pi) - \tfrac{1}{2} \alpha^{-2} \xi^2(t_i/\beta) + \ln \xi'(t_i/\beta) \}.$$

Then

$$-\frac{\alpha^3}{n} \frac{\partial L}{\partial \alpha} = \alpha^2 - \frac{1}{n} \sum_{i=1}^{n} \xi^2(t_i/\beta) \qquad (2.2.1)$$

$$\frac{\partial L}{\partial \beta} = -\frac{n}{\beta} + (\alpha\beta)^{-2} \sum_{i=1}^{n} t_i \xi(t_i/\beta) \xi'(t_i/\beta) - \frac{1}{\beta^2} \sum_{i=1}^{n} \frac{t_i \xi''(t_i/\beta)}{\xi'(t_i/\beta)} .$$

But we notice that

$$\xi(t) = \sqrt{t} - \frac{1}{\sqrt{t}} \qquad\qquad \xi^2(t) = t + \frac{1}{t} - 2$$

$$\xi'(t) = \frac{t^{\frac{1}{2}} + t^{-\frac{1}{2}}}{2t} = \frac{1}{2\xi(t)} (1 - 1/t^2) \qquad (2.2.2)$$

$$\frac{t\xi''(t)}{\xi'(t)} = -1 + \frac{1}{2} \frac{t-1}{t+1} = -\frac{1}{2} - \frac{1}{t+1} .$$

Hence by substitution we have

$$\frac{\partial L}{\partial \beta} = \frac{-n}{2\beta} + \frac{n}{2\alpha^2\beta} (\frac{s}{\beta} - \frac{\beta}{r}) + \frac{n}{K(\beta)} \qquad (2.2.3)$$

$$\frac{2\alpha^2\beta}{n} \frac{\partial L}{\partial \beta} = -\alpha^2 + (\frac{s}{\beta} - \frac{\beta}{r}) + \frac{2\alpha^2\beta}{K(\beta)} . \qquad (2.2.4)$$

By equating (2.2.1) to zero we then have the equation

$$\alpha^2 = \frac{s}{\beta} + \frac{\beta}{r} - 2 \qquad (2.3.1)$$

and by equating (2.2.4) to zero

$$\alpha^2 = \frac{s}{\beta} - \frac{\beta}{r} + \frac{2\beta\alpha^2}{K(\beta)} . \qquad (2.3.2)$$

Equating (2.3.1) and (2.3.2) and simplifying we obtain

$$\frac{\beta}{r} = 1 + \frac{\beta\alpha^2}{K(\beta)} .$$

Substituting from (2.3.1) for α^2 and simplifying we have that the maximum likelihood estimation of β is the solution of $g(x) = 0$ for $0 < x < \infty$ (presuming for the present it is unique). Now we will argue that $\hat{\beta}$ is unique and $r < \hat{\beta} < s$. Note that $g(0) = r[s+K(0)] = r(s+r) > 0$. We check that $g(x) \to -\infty$ as $x \to \infty$, by seeing that $\frac{K(x)}{x} \to 1$ and $[x-K(x)] \to -s$ as $x \to \infty$. Then

$$\frac{g(x)}{x} = x - K(x) + r\frac{K(x)}{x} - 2r + \frac{sr}{x}$$

so $g(x)/x \to -(s+r)$ as $x \to \infty$. Now

$$g'(x) = (x-r)[1-K'(x)] + x - r - K(x) \qquad (2.4.1)$$

and one sees that

$$K'(x) = K^2(x) \frac{1}{n} \Sigma(x+t_i)^{-2}, \qquad (2.4.2)$$

and we know, since $(E|X|^\nu)^{1/\nu}$ is nondecreasing in ν for any r.v. X, we have $K'(x) > 1$. Therefore, since $x - K(x)$ is decreasing we see that $K(x) > x-r$ for $x > 0$ and we have $g'(x) < 0$ for $x > r$ with at most one change of sign. Thus we have shown $\hat{\beta}$ unique now calculate

$$g(r) = r(s-r), \qquad g(s) = (s-r)[s-K(s)].$$

Again by the argument above we know that $s > r$, $\therefore$ $g(r) > 0$. Thus there exists a unique solution to $g(x) = 0$, say $\hat{\beta}$, and $\hat{\beta} > r$. But $g(s) < 0$ iff

$$\frac{1}{s} > \frac{1}{K(s)} \quad \text{iff} \quad 1 > \frac{1}{n}\sum\frac{s}{s+t_i} = 1 - \frac{1}{n}\sum\frac{t_i}{s+t_i} \quad (2.5)$$

therefore $g(s) < 0$. Thus the unique solution $\hat{\beta}$ is such that

$$s > \hat{\beta} > r.$$

Now that we know the unique solution of $g(x) = 0$ exists we check that it is indeed the maximum likelihood solution. Since by substituting (2.3.1) into (2.2.3) we have

$$\frac{1}{n}\frac{\partial L}{\partial \beta} = \frac{(r-\beta)}{rs+\beta^2-2\beta r} + \frac{1}{K(\beta)}$$

it is sufficient to check that $\left.\frac{\partial L}{\partial \beta}\right|_{\beta=r} > 0$, $\left.\frac{\partial L}{\partial \beta}\right|_{\beta=s} < 0$. By substituting into the above we find the first is $1/K(r) > 0$ and the second is $-1/s + 1/K(s) < 0$ by Equation (4.5).$||$

We now present two methods of finding $\hat{\beta}$.

Method I: If the data satisfies the inequality

$$2s < 3r + \min(t_1,\ldots,t_n) \quad (2.6)$$

then the Newton iteration procedure

$$\beta_{n+1} = \beta_n - \frac{g(\beta_n)}{g'(\beta_n)} \quad n=0,1,\ldots$$

will converge to $\hat{\beta}$ for all initial points $r < \beta_0 < s$, where the function g was defined in (2.1.3).

Proof. To assure ourselves that the Newton method will converge it is sufficient that g', g'' do not vanish for $r < x < s$ since they are continuous.

By referring to the formula for $g'(x)$ in (2.4.1) we see that for $r < x < s$, the first term is negative since $K'(x) > 1$. The second term is $x-r-K(x)$ which for $x = r$ is negative and decreases for $x > r$. Thus $g'(x) < 0$ for $r < x < s$ and cannot vanish in that interval.

We now show that $g''(x) > 0$ for $r < x < s$. Note

$$g''(x) = 2[1-K''(x)] - (x-r)K''(x),$$

and by taking derivatives of Equation (2.4.2) we note

$$\frac{K''(x)}{2K^3(x)} = [\frac{1}{n} \Sigma(x+t_i)^{-2}]^2 - \frac{1}{nK(x)} \Sigma(x+t_i)^{-3}.$$

Thus $g''(x) > 0$ for $r < x < s$ iff

$$\frac{1-K'(x)}{K^3(x)} < \frac{(x-r)K''(x)}{2K^3(x)}. \tag{2.6.1}$$

But the left-hand side of (2.6.1) is equal to

$$[\frac{1}{n} \Sigma(x+t_i)^{-1}]^3 - [\frac{1}{n} \Sigma(x+t_i)^{-1}][\frac{1}{n} \Sigma(x+t_i)^{-2}].$$

Multiplying both sides of (2.6.1) by $(x-r)^3$ and letting X (in this argument only) denote the random variable which takes the values $(x-r)/(x+t_i)$ for $i=1,\ldots,n$ with equal probability, we find that (2.6.1) is equivalent with

$$EXEX^2 - (EX)^3 > EXEX^3 - (EX^2)^2. \tag{2.6.2}$$

Recalling that $\ln EX^\nu$ is convex in ν we note that each side of the inequality (2.6.2) is positive.

Since $EX > 0$, we can divide (2.6.2) by it and obtain

$$EX^2 - (EX)^2 \geq EX^3 - \frac{(EX^2)^2}{EX} .$$

But because $(EX^\nu)^{1/\nu}$ is nondecreasing in ν we know that $(EX^2)^2 \geq (EX)^4$. Hence

$$EX^3 - \frac{(EX^2)^2}{EX} \leq EX^3 - (EX)^3$$

and it is sufficient for (2.6.2) to show that $EX^2 - (EX)^2 \geq EX^3 - (EX)^3$, or equivalently to show $EX^2(1-X) \geq (EX)^2(1-EX)$. But for any convex ψ we know $E\psi(X) \geq \psi(EX)$. Thus $\psi(x) = x^2 - x^3$ is convex for $x < 1/3$. Therefore, a sufficient condition that $g''(x)$ not vanish on $r < x < s$ is

$$\frac{x-r}{x+t_i} < \frac{1}{3} \quad \text{for all} \quad i=1,\ldots,n. \tag{2.6.3}$$

One can check that (2.6) implies (2.6.3) in the interval needed.$\|$

Method II: If the data satisfies

$$2r > s, \tag{2.7}$$

then for $H = \frac{1}{2} K$, where K was defined in Equation (2.1.2), set

$$A(x) = r + H(x) - \sqrt{H^2(x)-r(s-r)} \tag{2.7.0.1}$$

and for all initial points $r < \beta_0 < s$, as $n \to \infty$ the iteration $A^{(n)}(\beta_0)$ converges to $\hat{\beta}$.

Proof. Solve $g(x) = 0$ in the form

$$x^2 - 2x[r+H(x)] + r[s+2H(x)] = 0$$

by considering $H(x)$ as a constant. Then using the quadratic formula and selecting the appropriate root we find the solution is $A(x)$ as defined in (2.7.0.1).

In order to assure ourselves that the values $A(x)$ are real for $x > r$ we need to show

$$H(x) > \sqrt{r(s-r)} \; . \qquad (2.7.1)$$

A sufficient condition that this is true is (2.7) as we now show. At $x = r$ (2.7.1) is equivalent with

$$1/\sqrt{r(s-r)} > \frac{2}{n} \sum_1^n \frac{1}{r+t_i} \; ,$$

but by the relationship of harmonic and arithmetic means we know

$$\frac{2}{r+t_i} < \frac{1}{2}(\frac{1}{r} + \frac{1}{t_i})$$

and hence

$$\frac{1}{H(r)} = \frac{1}{n} \sum \frac{2}{r+t_i} < \frac{1}{2}(\frac{1}{r} + \frac{1}{r}) = \frac{1}{r} \; .$$

Thus by assumption (2.7), the inequality (2.7.1) will be satisfied since we know $H > r$ and $H' > 0$ for all $x > 0$.

One notes that

$$A' = H' \left[1 - \frac{H}{\sqrt{H^2 - r(s-r)}} \right]$$

and by the noted properties of H we see that $A' < 0$. Hence A is monotone decreasing. Thus if we can show for $r < \beta_0 < s$ we have $A^{(2)}(\beta_0)$ bounded (here (2) indicates composition with itself) then by

the uniqueness of the maximum likelihood estimate (which we have seen previously) and the completeness of the real line we have $A^{(n)}(\beta_0) \to \hat{\beta}$.

From the definition, using the first two terms of the binomial expansion we see,

$$r + \frac{r(s-r)}{2H(\beta_0)} \le A(\beta_0) \le r + H(\beta_0). \qquad (2.7.2)$$

Applying A again we find

$$A[r + \frac{r(s-r)}{2H(\beta_0)}] \ge A^{(2)}(\beta_0) \ge A[r + H(\beta_0)].$$

But by (2.7.2) we have

$$r + \frac{r(s-r)}{2H[r+H(\beta_0)]} \le A^{(2)}(\beta_0) \le r + H[r + \frac{r(s-r)}{2H(\beta_0)}]. \qquad (2.7.3)$$

For this argument let Y be the random variable taking the values t_i for $i=1,\ldots,n$ with probability $1/n$, now since the function of y defined by $2/(x+y)$ is convex in y and $EY = s$ we have by Jensen's inequality

$$\frac{1}{H(x)} = E(\frac{2}{x+Y}) \ge \frac{1}{x+EY} = \frac{1}{x+s} .$$

Applying this inequality to the right-hand side of (2.7.3):

$$A^{(2)}(\beta_0) \le 2r + s + \frac{r(s-r)}{2H(\beta_0)} .$$

Using (2.7.1) we obtain a condition known to be sufficient for convergence,

$$r < A^{(2)}(\beta_0) < s + 2r + \tfrac{1}{2}\sqrt{r(s-r)} .||$$

We claim that the conditions (2.6) and (2.7) are not stringent in practice and will nearly always be satisfied.

3. The Mean Mean

Since all iteration methods are functionally related to how close the initial guess is to the answer sought, we take $\beta_0 = \tilde{\beta}$ where

$$\tilde{\beta} = (SR)^{\frac{1}{2}} \tag{3.1}$$

where S and R, now considered as random variables and written in the upper case, were defined in (2.1.1). We choose as the initial estimate for the median life β the geometric mean of the harmonic and arithmetic means of the sample lives in the data: we call this estimate $\tilde{\beta}$ the *mean mean*. We assert that the great utility of this estimate which we shall demonstrate subsequently can be accounted for in the following comments.

Theorem 3.1. $\tilde{\beta}$ is a consistent estimate for β.

Proof. By (1.3) and the strong law of large numbers, with probability one

$$S_n = \frac{1}{n} \sum_1^n T_i \to ET = \beta(1 + \frac{\alpha^2}{2}) \quad \text{as} \quad n \to \infty.$$

But also $\frac{1}{R_n} = \frac{1}{n} \sum_1^n (\frac{1}{T_i})$ and $\frac{1}{T_i}$, by Theorem 1.1 has the distribution $F(\alpha, \frac{1}{\beta})$ so that with probability one

$$\frac{1}{R_n} \to E(\frac{1}{T}) = \frac{1}{\beta}(1 + \frac{\alpha^2}{2}) \quad \text{as} \quad n \to \infty.$$

Thus clearly $(\tilde{\beta}_n)^2 = S_n R_n \to \beta^2$ as $n \to \infty$ with probability one.||

We now prove

Theorem 3.2. If $\tilde{\beta}$ is a fixed point of $H = \frac{1}{2}K$, then $\tilde{\beta} = \hat{\beta}$.

Proof. Using the hypothesis that $H(\tilde{\beta}) = \tilde{\beta}$ we find from (2.1.3)

$$g(\tilde{\beta}) = (\tilde{\beta})^2 - 2\tilde{\beta}r - 2(\tilde{\beta})^2 + (\tilde{\beta})^2 + 2r\tilde{\beta} = 0$$

but by definition $g(\tilde{\beta}) = 0$ implies $\hat{\beta} = \tilde{\beta}$, since $\hat{\beta}$ is the unique solution of $g(x) = 0$ for $r < x < s.||$

Theorem 3.3. If $\hat{\beta}$ is a fixed point of $H = \dfrac{K}{2}$, then $\hat{\beta} = \tilde{\beta}$.

Proof. By definition $\hat{\beta}$ is a fixed point of A, hence

$$\hat{\beta} = A(\hat{\beta}) = r + H(\hat{\beta}) - \sqrt{H^2(\hat{\beta}) - r(s-r)}$$

but by hypothesis $H(\hat{\beta}) = \hat{\beta}$ so that $\hat{\beta} = r + \hat{\beta} - \sqrt{(\hat{\beta})^2 - r(s-r)}$ and solving for $\hat{\beta}$ shows that $\hat{\beta} = \sqrt{rs}.||$

If in practice the conditions should obtain for which the mean mean is a fixed point of H, within a reasonable approximation, it would follow from the preceding two results that within that degree of approximation we could take the mean mean as the maximum likelihood estimate of β.

We now present some sufficient conditions that the mean mean of a set of numbers be a fixed point of H.

Theorem 3.4. If $t_1, \ldots, t_{2k}$ satisfy the relations

$$t_{2i-1} = \beta\tau_i, \qquad t_{2i} = \beta/\tau_i \qquad \text{for} \qquad i=1,\ldots,k \tag{3.2}$$

or the relations

$$t_{2i-1} = \frac{\tau_i}{\beta}, \qquad t_{2i} = \frac{\beta}{\tau_i} \qquad \text{for} \qquad i=1,\ldots,k \tag{3.3}$$

where $\beta, \tau_1, \ldots, \tau_k$ are any set of positive numbers, then the mean mean $\tilde{\beta}$ from these samples satisfies exactly the identity $H(\tilde{\beta}) = \hat{\beta}$.

Proof. By definition we must show that $1 = 2\tilde{\beta}/K(\tilde{\beta})$ but note that in the case (3.2) that $\tilde{\beta} = \beta$ since

$$\frac{s}{\beta} = \frac{1}{2k}\left[\sum_{1}^{k}\tau_i + \sum_{1}^{k}\frac{1}{\tau_i}\right] = \frac{1}{r\beta}$$

$$\therefore \quad \frac{\beta}{K(\beta)} = \frac{1}{2k}\sum_{1}^{2k}\frac{\beta}{\beta+t_i} = \frac{1}{2k}\left[\sum_{1}^{k}\frac{1}{1+\tau_i} + \sum_{1}^{k}\frac{1}{1+\tau_i^{-1}}\right] = \frac{1}{2} .$$

Now in the case (3.3) we see that $\tilde{\beta} = 1$ and the remainder of the argument follows from that just given with $\beta = 1$.

Corollary 3.4. The conclusion is the same if

in case (3.2) we add $t_{2k+1} = \beta$, and

in case (3.3) we add $\tau_{2k+1} = 1$.

We now claim that any sample from a distribution in $\mathcal{G}$ tends to act like the sets of numbers described in (3.2) and (3.3). If $T_1,\ldots,T_{2k}$ are independent random variables each with the distribution $F(\alpha,\beta)$ in $\mathcal{G}$, then by Theorem 1.1 we see that $\frac{\beta}{T_1},\ldots,\frac{\beta}{T_k}$ each has the distribution $F(\alpha,1)$ which is exactly the same as the distribution of $\frac{T_{k+1}}{\beta},\ldots,\frac{T_{2k}}{\beta}$. Thus we see that a sample tends to satisfy (3.2).

But on the other hand, if $T_1,\ldots,T_k,\frac{1}{T_{k+1}},\ldots,\frac{1}{T_{2k}}$ are all independent and identically distributed by $F(\alpha,1)$, then $\beta T_1,\ldots,\beta T_k$ have distribution $F(\alpha,\beta)$ while $\frac{\beta}{T_{k+1}},\ldots,\frac{\beta}{T_{2k}}$ also have the same common distribution. Again we see that a sample tends to satisfy (3.3). Thus we see why $\tilde{\beta}$ should be near a fixed point of H.

We now state

Theorem 3.5. If T has the distribution $F(\alpha,1)$ in $\mathcal{G}$, then

$$E(\frac{2}{1+T}) = 1 \quad \text{for all} \quad \alpha > 0.$$

Proof. By definition

$$E(\frac{1}{1+T}) = \int_0^1 \frac{1}{1+t} \, dF(t:\alpha,1) + \int_1^\infty dF(t:\alpha,1) - \int_1^\infty \frac{t}{1+t} \, dF(t:\alpha,1).$$

Making the change of variable $y = 1/t$ in the third integral and realizing that $F(\frac{1}{s} : \alpha,1) = 1 - F(s:\alpha,1)$ we find that the first and third terms cancel. Hence

$$E(\frac{1}{1+T}) = 1 - \mathfrak{R}(0) = \frac{1}{2} \cdot ||$$

We use this theorem to obtain a result which says that under any conditions, for a sufficiently large sample, $\tilde{\beta}$ will be nearly a fixed point of H. This is because $\tilde{\beta}/H(\tilde{\beta})$ approaches unity as the sample size approaches infinity. Note that this result is not obtainable directly from Slutsky's theorem, given e.g., p. 255, Ref. [4], since H itself is a random function depending upon the sample size despite our suppression of this fact with our notation.

Theorem 3.6. The random variable $\tilde{\beta}_n/H(\tilde{\beta}_n)$ converges in probability to 1 as $n \to \infty$.

Proof. For this argument let us set

$$S_n(x) = \frac{2}{n} \sum_1^n \frac{x}{x+T_i}$$

where each T_i has distribution $F(\alpha,\beta)$. Of course we see that $S_n(\tilde{\beta}_n) = \tilde{\beta}_n/H(\tilde{\beta}_n)$. It is sufficient to show, by Theorem 3.5, that

$$S_n(\tilde{\beta}_n) \xrightarrow{p} S(\beta) = E(\frac{\beta}{\beta+T}).$$

Now

$$P[|S_n(\tilde{\beta}_n)-S(\beta)| \geq \epsilon] \leq P[|S_n(\tilde{\beta}_n)-S(\tilde{\beta}_n)| \geq \epsilon] + P[|S(\tilde{\beta}_n)-S(\beta)| \geq \epsilon].$$

Since

$$E(\frac{x}{x+T})^2 = \int_0^\infty \frac{1}{(1+t)^2} \, dF(t:\alpha,\beta/x) \leq 1$$

we conclude that $\mathrm{var}(S_n(x)) \leq \dfrac{C_1}{n}$ uniformly in x.

By Chebyscheff's inequality as $n \to \infty$

$$P[|S_n(\tilde{\beta}_n)-S(\tilde{\beta}_n)| \geq \epsilon] \leq \frac{\mathrm{var}\, S_n(\tilde{\beta}_n)}{\epsilon^2} \leq \frac{C_1}{n\epsilon^2} \to 0.$$

But, of course, the second term approaches zero by the consistency of $\tilde{\beta}_n$ for β. ||

The asymptotic behavior of maximum likelihood estimates such as $\hat{\beta}$ is well know. We now prove an important result concerning the asymptotic distribution of the mean mean $\tilde{\beta}$.

<u>Theorem 3.7.</u> For n sufficiently large the estimator $\tilde{\beta}_n$ is asymptotically distributed by $F(\dfrac{\alpha\theta}{\sqrt{n}},\beta)$ in $\mathcal{G}$ where

$$\theta^2 = (1 + \frac{3\alpha^2}{4})/(1 + \frac{\alpha^2}{2})^2. \tag{3.3.1}$$

Proof. By definition $\tilde{\beta}_n = \beta\sqrt{Y_n}$ where $Y_n = (\sum_1^n T_i)/(\sum_1^n T_i^{-1})$ with T_i independently and identically distributed by $F(\alpha,1)$. Let

$$U_i = Z_i\sqrt{1 + \frac{\alpha^2}{4} Z_i^2} \qquad \text{for} \qquad i=1,\ldots,n \tag{3.4}$$

where Z_i are independent $\mathfrak{N}(0,1)$ variates. By (1.5) we may write

$$Y_n = (1 + \alpha X_n)/(1 - \alpha X_n) \tag{3.5}$$

where

$$\sqrt{n} \ X_n = \frac{\sum_1^n U_i}{\sqrt{n} \ (1 + \frac{\alpha^2}{2n} \sum_1^n Z_i^2)} \ . \tag{3.6}$$

To prove the result it is sufficient to show that as $n \to \infty$,

$$\frac{\sqrt{n}}{\alpha} \ \xi(\frac{\tilde{\beta}_n}{\beta}) = \frac{\sqrt{n}}{\alpha} \ [(Y_n)^{\frac{1}{4}} - (Y_n)^{-\frac{1}{4}}] \tag{3.7}$$

converges in distribution to the $\mathfrak{N}(0,\theta^2)$ law.

Substituting (3.5) into (3.7) and applying the binomial expansion, which is valid for n sufficiently large we find

$$\frac{\sqrt{n}}{\alpha} \ \xi(\frac{\tilde{\beta}_n}{\beta}) = \frac{\sqrt{n}}{\alpha} \ [\alpha X_n + c_3 \alpha^3 X_n^3 + \cdots]$$

where $c_k = 2 \sum_{j=0}^{k} \binom{\frac{1}{4}}{k-j}\binom{-\frac{1}{4}}{j}(-1)^j$ for k odd. It follows from (3.6) since $EU_i = 0$, $var(U_i) = 1 + \frac{3\alpha^2}{4}$, that $E(\sqrt{n} \ X_n) \to 0$, $var(\sqrt{n} \ X_n) \to \theta^2$, and $\sqrt{n} \ X_n$ converges in distribution to a normal law.||

Combining this result with Theorem 1.1 we have the immediate

Corollary 3.8. For n sufficiently large, with θ defined in (3.3.1),

$$E\tilde{\beta}_n = \beta[1 + \frac{(\alpha\theta)^2}{2n}]$$

and

$$var(\tilde{\beta}_n) = \frac{(\alpha\theta\beta)^2}{n} \ [1 + \frac{5\alpha^2\theta^2}{4n}].$$

Having examined the behavior of $\tilde{\beta}_n$ for n sufficiently large, in particular its mean and variance, we now fix n and determine the mean and variance of $\tilde{\beta}_n$ for α small. Note that for α small we have

$$\theta^2 = 1 - \frac{\alpha^2}{4} + O(\alpha^4),$$

which is nearly unity.

We also state

<u>Theorem 3.8</u>. The estimate $\tilde{\beta}_n$ is for each fixed n biased by a factor b_n which, for α sufficiently small, is given by

$$b_n = 1 + \frac{\alpha^2}{2n} + 0(\alpha^4)$$

with

$$var(\tilde{\beta}_n) = \frac{(\alpha\beta)^2}{n} + 0(\alpha^4).$$

Proof. The proof follows using similar techniques and notation to that utilized in Theorem 3.7.

For n fixed, any real h and any random variables U_i $i=1,\dots,n$ we define $<U_i^h> = \frac{1}{n} \sum_{i=1}^{n} U_i^h$. The bias factor of $\tilde{\beta}_n$ is $b_n = E\sqrt{Y_n}$, where Y_n was defined in (3.4). From (3.5) we have

$$X_n = <U_i> [1 + \frac{\alpha^2}{2} <Z_i>]^{-1} \tag{3.8}$$

where here the U_i are as defined in (3.4). On the set $B_\alpha = \bigcap_{i=1}^{n} [|Z_i| < \frac{\sqrt{2}}{\alpha}]$, we may apply the binomial expansion to the conditional random variable

$$<U_i>|B_\alpha = \sum_{j=0}^{\infty} \binom{\frac{1}{2}}{j} (\frac{\alpha}{2})^{2j} <Z_i^{2j+1}>.$$

Since $B_\alpha \subset [<Z_i^2> < \frac{2}{\alpha^2}]$ by using (3.8) we have, after applying the binomial expansion and simplifying,

$$X_n|B_\alpha = \sum_{k=0}^{\infty} \alpha^{2k} \sum_{j=0}^{k} c_{kj} <Z_i^{2j+1}> <Z_i^2>^{k-j}$$

where for notational simplicity we set $c_{kj} = \binom{\frac{1}{2}}{j}\binom{-\frac{1}{2}}{k-j} 4^{-j} 2^{j-k}$. From the definition (3.5)

$$\sqrt{Y_n} = (1 + \alpha X_n)^{\frac{1}{2}} (1 - \alpha X_n)^{-\frac{1}{2}},$$

again since $B_\alpha \subset [|X_n| < \frac{1}{\alpha}]$, by expanding and simplifying we obtain

$$\sqrt{Y_n} \, |B_\alpha = \sum_{j=0}^{\infty} \binom{\frac{1}{2}}{j} \alpha^j X_n^j \sum_{i=0}^{\infty} \binom{-\frac{1}{2}}{i} (-\alpha)^i X_n^i.$$

By rearranging we have

$$\sqrt{Y_n} \, |B_\alpha = \sum_{k=0}^{\infty} \alpha^k c_k X_n^k$$

where

$$c_k = \sum_{i=0}^{k} \binom{-\frac{1}{2}}{i} \binom{\frac{1}{2}}{k-i} (-1)^i.$$

By definition

$$b_n = \sum_{i=0}^{\infty} \alpha^k c_k E[X_n \{B_\alpha\}]^k + E[\sqrt{Y_n} \{B_\alpha^c\}] \tag{3.9}$$

where $\{B\}$ is the indicator function of that set B. Note

$$E[\sqrt{Y_n} \{B_\alpha^c\}] \leq (EY_n)[1 - P(B_\alpha)]$$

and $P(B_\alpha) = [\mathfrak{N}(\frac{\sqrt{2}}{\alpha}) - \mathfrak{N}(\frac{-\sqrt{2}}{\alpha})]^n \to 1$ as $\alpha \to 0$.

Using (3.7)

$$X_n | B_\alpha = <Z_i> + \alpha^2 c_{10} <Z_i^2><Z_i> + \alpha^2 c_{11} <Z_i^3> + 0(\alpha^4)$$

$$[X_n | B_\alpha]^2 = <Z_i>^2 + 2\alpha^2 c_{10} <Z_i^2><Z_i>^2 + 2\alpha^2 c_{11} <Z_i^3><Z_i> + 0(\alpha^4)$$

$$[X_n | B_\alpha]^3 = <Z_i>^3 + 0(\alpha^2)$$

and for $k \geq 4$

$$[X_n | B_\alpha]^k = 0(\alpha^{k-1}).$$

Taking the expectation and realizing that $Z_1 | [|Z_1| < \frac{\sqrt{2}}{\alpha}]$ is a symmetric random variable we see

$$E[X_n \{ B_\alpha \}] = 0(\alpha^4)$$

$$E[X_n \{ B_\alpha \}]^2 = \frac{1}{n} + 0(\alpha^2)$$

$$E[X_n \{ B_\alpha \}]^3 = 0(\alpha^2).$$

Since $c_0 = c_1 = 1$, $c_2 = c_3 = \frac{1}{2}$ we conclude from (3.8) that (3.4) holds.

Consider the variance of $\tilde{\beta}_n$. We know

$$var(\tilde{\beta}_n) = \beta^2(EY_n - b_n^2).$$

Proceeding as before have

$$Y_n | B_\alpha = 1 + 2\alpha X_n + 2\alpha^2 X_n^2 + 2\alpha^3 X_n^3 + \cdots$$

$$EY_n \{ B_\alpha \} = 1 + \frac{2\alpha^2}{n} + 0(\alpha^4)$$

and by squaring (3.4) and subtracting we obtain (3.5).||

This completes our theoretical study of the behavior of the mean mean. In summary, we have seen that $\tilde{\beta}_n$ is a consistent estimator of β which for α small, say less than $1/2$, is almost unbiased while having a variance of approximately $(\alpha\beta)^2/n$. Moreover, we claim that under this condition, which we shall later empirically verify, $\tilde{\beta}_n$ is virtually the maximum likelihood estimator whose optimal properties are well known.

4. Numerical Examples and Comparisons

Let us first perform an empirical comparison of the two iterative procedures for obtaining the maximum likelihood estimator $\hat{\beta}$, and check to see if the known consistency of this estimate is of practical significance. We shall obtain a set of observations from the random variable defined by (1.5) by selecting α,β and then using a library machine program for the IBM 360 to generate the pseudo-random normal observations. We know that the computed random variable has a distribution in $\mathcal{G}$.

For given $t_1,\ldots,t_n$ generated by this method we compute the sample mean mean $\tilde{\beta} = \sqrt{sr}$, where s and r are defined in (2.1.1). (As we know $\tilde{\beta}$ is itself a consistent estimate of β.) We stop the iteration when the successive iterates agree within a prescribed ε.

The results of this simulation are presented in Table I. Method II, which computationally is the simpler, appears to work as well for $\alpha \leq .5$ as the well-known Newton procedure, however, it does not work at all for values of α as large as 2.

The important point is that for $\alpha < .5$ there is no real need to compute anything but $\tilde{\beta}$ since both estimates agreed to within three significant digits. It is doubtful that any further iteration by either method would really improve the accuracy of the estimate.

The computational simplicity of the estimate $\tilde{\beta}$, as well as the advantages of its asymptotic distribution being in the class $\mathcal{G}$, not to mention the properties which were set out in the last paragraph of the proceeding section, can all be utilized if the range of α which is

encountered in practice is no larger than say 1/2. Thus in fatigue

studies it would be desirable that the appropriate value of α were

within this range.

Table I

Comparison of methods to find MLE within ϵ for $\beta = 100$, $\epsilon = .001$

sample size n	$\tilde{\beta}$	$\hat{\beta}$	$\hat{\alpha}$	number of iterations for convergence within ϵ	
				I	II
$\alpha = .5$					
5	105.119	105.117	.184	2	2
10	99.611	99.580	.358	2	3
25	87.279	87.281	.602	2	2
50	110.117	110.125	.497	2	2
100	101.201	101.194	.590	2	2
250	99.676	99.702	.496	2	3
500	100.215	100.207	.476	2	2
$\alpha = 1$					
5	79.180	79.200	.363	2	2
10	67.302	68.584	.986	3	8
25	113.710	116.176	1.095	3	10
50	86.470	86.225	1.091	2	7
100	113.666	114.202	.866	2	6
250	105.968	106.156	.964	2	6
500	94.973	95.042	1.016	2	5
$\alpha = 2$					
5	305.983	310.994	1.748	3	*
10	48.948	52.314	1.654	3	*
25	125.3??	131.245	1.796	3	*
50	96.638	99.622	1.652	3	*
100	87.675	88.032	2.110	2	*
250	99.848	99.634	2.030	2	*
500	96.640	96.696	2.127	2	*

*Method II failed to be applicable since for some point β_n in the iteration we have $K(\beta_n) < r(s-r)$ and β_{n+1} is no longer real.

We now confront this family $\mathcal{G}$ of distributions with some actual fatigue data. We choose some extensive data on the fatigue life of 6061-T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second. Some of this data has been reported earlier in [2].

For a maximum stress per cycle of 31,000 psi we give the 101 observations of lifetimes in cycles $X10^{-3}$.

SAMPLE SIZE = 101

70	90	96	97	99	100	103	104
104	105	107	108	108	108	109	109
112	112	113	114	114	114	116	119
120	120	120	121	121	123	124	124
124	124	124	128	128	129	129	130
130	130	131	131	131	131	131	132
132	132	133	134	134	134	134	134
136	136	137	138	138	138	139	139
141	141	142	142	142	142	142	142
144	144	145	146	148	148	149	151
151	152	155	156	157	157	157	157
158	159	162	163	163	164	166	166
168	170	174	196	212			

Setting $\epsilon = .0001$ we find using Method I, starting with $\tilde{\beta} = 131.819454$ that in three iterations

$$\hat{\beta} = 131.81903$$

$$\hat{\alpha} = .17037302,$$

while with Method II, with the same value of $\tilde{\beta}$ as initial guess, we find that in two iterations

$$\hat{\beta} = 131.81895$$

$$\hat{\alpha} = .17037022.$$

From the graphical comparison of the fitted dis ribution to the empiric cumulative in Figure 1, we see that $\mathcal{F}$ provides an adequate explanation.
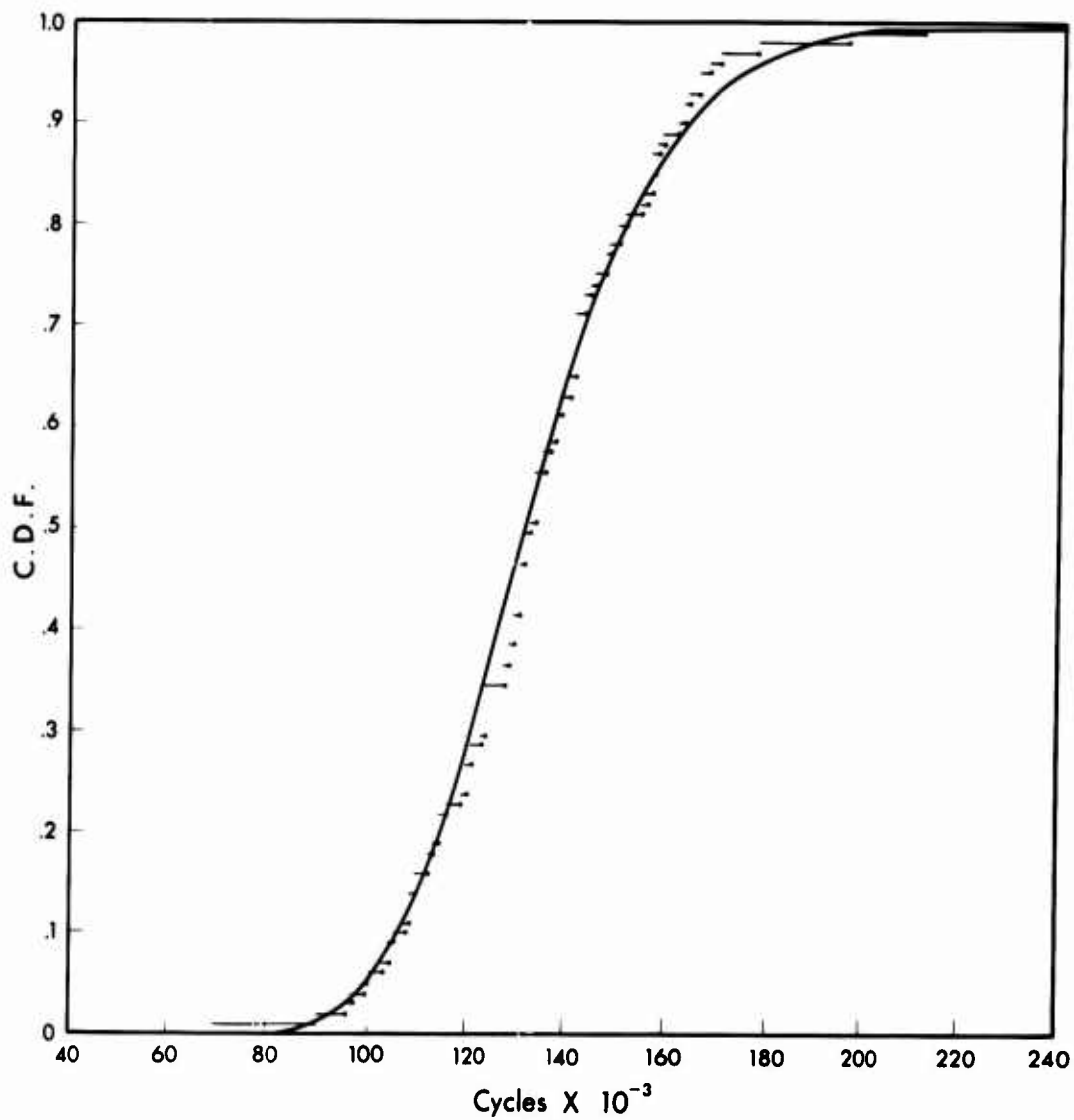

Figure 1

The empiric cumulative and the distribution $F(\hat{\alpha},\hat{\beta})$ for fatigue life at a stress of 31,000 psi.
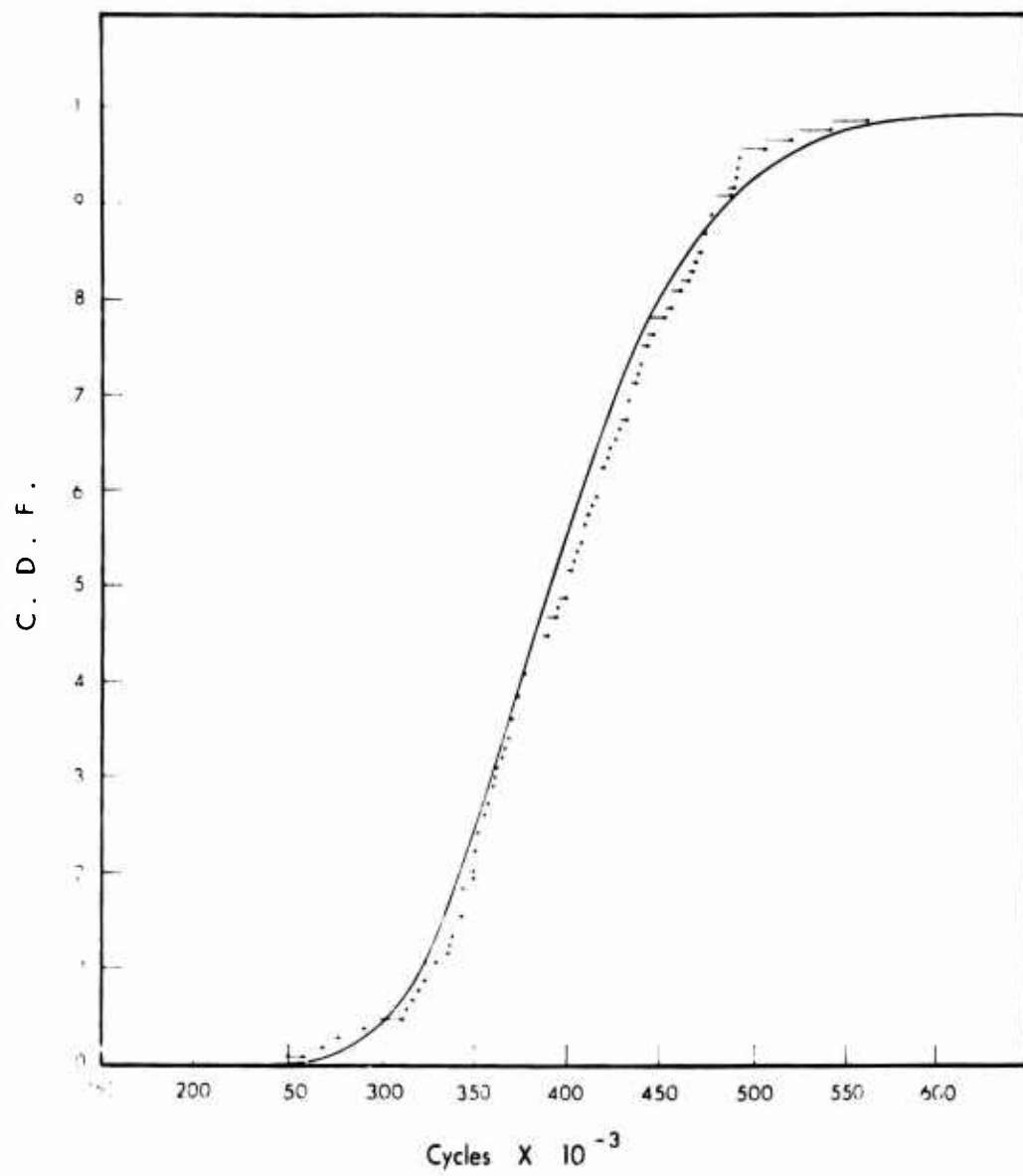
Figure 2

The empiric cumulative and the distribution $F(\hat{\alpha}, \hat{\beta})$ for fatigue life at a stress of 26,000 psi.

For a maximum stress of 26,000 psi we give the 102 observations of lifetimes in cycles $X10^{-3}$

SAMPLE SIZE = 102

233	258	268	276	290	310	312	315
318	321	321	329	335	336	338	338
342	342	342	344	349	350	350	351
351	352	352	356	358	358	360	362
363	366	367	370	370	372	372	374
375	376	379	379	380	382	389	389
395	396	400	400	400	403	404	406
408	408	410	412	414	416	416	416
420	422	423	426	428	432	432	433
433	437	438	439	439	443	445	445
452	456	456	460	464	466	468	470
470	473	474	476	476	486	488	489
490	491	503	517	540	560		

Again choosing $\varepsilon = .0001$ we find using Method I starting with $\tilde{\beta} = 392.765189$, that in two iterations

$$\hat{\beta} = 392.76367$$

$$\hat{\alpha} = .16141957$$

and that for Method II in two iterations

$$\hat{\beta} = 392.76416$$

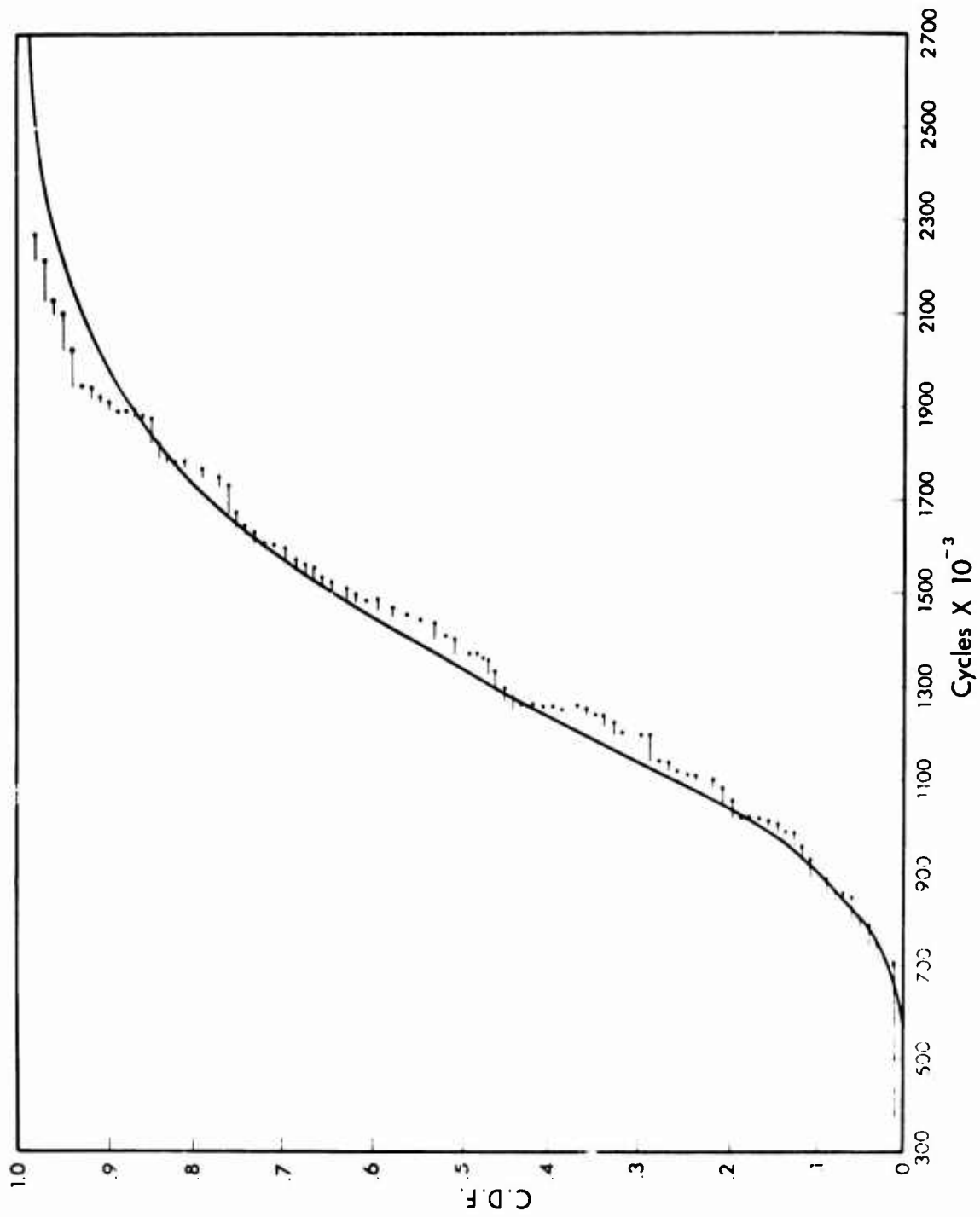$$\hat{\alpha} = .1614195.$$

See Figure 2 for a graphical comparison.

Figure 3

The empiric cumulative and the distribution $F(\hat{\alpha},\hat{\beta})$ for fatigue life at a stress of 21,000 psi.

For a maximum stress of 21,000 psi we have 101 observations of lifetimes in cycles $X10^{-3}$

SAMPLE SIZE = 101

370	706	716	746	785	797	844	855
858	886	886	930	960	988	990	1000
1010	1016	1018	1020	1055	1085	1102	1102
1108	1115	1120	1134	1140	1199	1200	1200
1203	1222	1235	1238	1252	1258	1262	1269
1270	1290	1293	1300	1310	1313	1315	1330
1355	1390	1416	1419	1420	1420	1450	1452
1475	1478	1481	1485	1502	1505	1513	1522
1522	1530	1540	1560	1567	1578	1594	1602
1604	1608	1630	1642	1674	1730	1750	1750
1763	1768	1781	1782	1792	1820	1868	1881
1890	1893	1895	1910	1923	1940	1945	2023
2100	2130	2215	2268	2440			

With $\varepsilon = .0001$, $\tilde{\beta} = 1336.56547$ we find that in fifty-one iterations Method I failed to converge yielding as values at that time

$$\hat{\beta} = 1336.3779$$

$$\hat{\alpha} = .31029648,$$

but that Method II converged in three iterations to

$$\hat{\beta} = 1336.3784$$

$$\hat{\alpha} = .31029648.$$

We believe that the extreme stringency of ε and round-off error in the machine arithmetic caused Method I to fail to converge and not a theoretical deficiency in the method itself.

We can tentatively conclude from this evidence that in fatigue applications the appropriate range of α is sufficiently small as to allow the use of $\tilde{\beta}$ as an estimate of the median β and thus one can utilize the properties which we have previously discussed.

References

[1] Birnbaum, Z. W., and Saunders, S. C. A New Family of Life
Distributions. Submitted for publication to *Journal of Applied
Statistics*.

[2] Birnbaum, Z. W., and Saunders, S. C. A Statistical Model for
Life Length of Material. *J. Amer. Statist. Assoc. 53*, 151-160,
1958.

[3] Buckland, W. R. Statistical Assessment of the Life Characteristic.
Griffin. London, 1964.

[4] Cramér, Harold Mathematical Methods of Statistics. Princeton
University Press. Princeton, 1951.

[5] Gumbel, E. J. Statistics of Extremes. Columbia University Press.
New York, 1958.

[6] Parzen, Emanuel On Models for the Probability of Fatigue Failure
of a Structure. Report 245, North Atlantic Treaty Organization
Advisory Group for Aeronautical Research and Development, 1959.